Libet's Experiment

Why It Matters and What It Means

Introduction 2

Section I: Libet's Experiment: A Brief Discussion 5

Section II: Follow-up Studies 11

Some Representative Studies 12

Discussion 13

Section III: Philosophers Respond to Libet 14

Preliminaries 14

Philosophers' Responses to the Libet Experiment 17

"How does the neuroscience of decision-making bear on our understanding of moral responsibility and free will" (2010) by Adina Roskies 17

"A Critique of Libet and Wegner's Argument Against Free Will" (2023) by Ferdinard Fosu-Blankson. 19

"Free Will and Neuroscience: Decision Times and the Point of No Return" (2019) by Alfred Mele. 22

"Evidentiating Free Will" (2014) by Nicholas Rescher. 25

"Does the brain 'initiate' freely willed processes? A philosophy of science critique of Libet-type experiments and their interpretation" (2013) by Hans Radder and Gerben Meynen ... 27

Some other recurring philosophical tropes 30

Libet confuses urges with intentions 30

Libet's results don't generalize because they concern actions that fail to engage the subject's rationality 32

It is a "phenomenological datum" that we have free will 33

Section IV. A constructive response to the Libet Experiment 34

Preliminaries 35

A Tentative Model 37

Conclusion 39

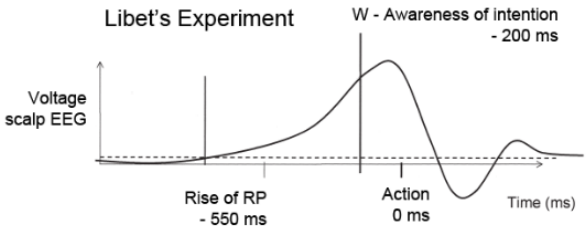
References 41

Introduction

In the early 1980s, a team of neuroscientists, headed by Benjamin Libet, conducted an experiment that had a striking outcome (Libet et al. 1983). The subject of the experiment simply had to press a button while making a mental note of when he decided to press the button. To this end, there was a clock in front of him with a rapidly moving millisecond hand to help the subject pinpoint the time at which he was first consciously aware of his intention to press the button. The subject was free to press the button whenever he wanted: he did not have to press the button when some external condition was met—when, for example, he heard a certain sound—it being entirely up to him when to press the button. Heat-sensitive pads were attached to the subjects cranium and connected via wires to a machine that measured electrical activity in the subject's brain activity on the basis of the resulting temperature-readings.

The results of the experiment were striking; they are embodied in Figures 1 and 2:

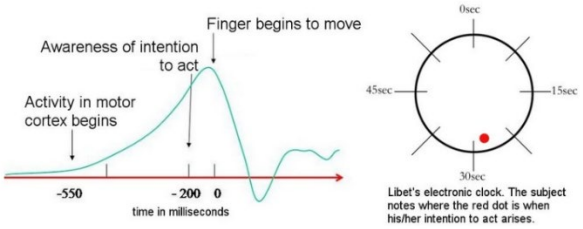
Figure 1



x

Taken from Libet 1993.

Figure 2



Taken from Løvstad et al. 2012.

Just as we would expect, this awareness slightly precedes the resulting finger-movement. But as the graph shows, the subject is conscious of his intention only after it has long been in the making. Notably, the subject always feels as though he "just came up with" the decision in question; that he "freely" chose to move his finger; and that his decision (i.e., the conscious event that he took to be his decision) was not the result of any previous mentation on his part (Libet 1993; Wegner 2004). His decision was *not*, so he felt, the result of deliberation; it was *not*, so he felt, the result of a process of deliberation or of any sort of psychological activity that might be dispositive of it. The subject described his decision as utterly "free" (Libet 1993, 1999; Wegner 2004).

This experiment was repeated, by Libet himself and by others, with hundreds of different subjects, always with the same results. The shape and length of the slope was always the same; as was the length of the gap between RP-initiation and W; as was the length of the gap between W and the corresponding finger-movement (Braun et al., 2021).

To expedite discussion, we will use the following abbreviations throughout this paper:

"RP": reaction potential.

"RPI": increase in reaction potential, i.e., the upward-sloping part of the graph.

"LE": Libet's experiment

"RLE": The results of Libet's experiment.

"W": The subject's first consciousness of his decision to move his finger.

In the present paper, we will delineate the philosophical implications of LE and in the process, we will consider what philosophers have said on the matter. By way of anticipation, here is what we will find. RLE are profoundly important; they do in fact warrant a revision of many conceits about the mind—specifically about agency and the role therein of the "self", whatever that is ---that are accepted both by philosophers and laymen. With few or no exceptions, philosophers who have written on the significance of LE have attempted to trivialize it (Schlosser 2013). Some have taken the position that the actual data generated by the experiment is spurious (Mele 2009, 2010). Some have taken the position that the brain-activity preceding W is not *causative* of W and, therefore, that W is indeed the operative intention (Radder & Meynen 2013; Rescher 2014; Mele 2019).

Most philosophers, if not all, have taken the position that, even if that data is legitimate, it doesn't warrant a revision of any of frameworks within which philosophers have traditionally attempted to understand the nature of human "freedom", whatever that might be, or, more generally, of human agency. According to some philosophers, LE only shows that, just as we already knew, the "compatibilist" conception of freedom is the right one (Roskies 2010). (Compatibilism: freedom is compatible with determinism and may even require it.) According to others, LE is relevant to questions concerning human agency *only* if compatibilism is correct and therefore has *no* such relevance, since, just as we already knew, compatibilism is false (Griffith 2013).

Significantly, philosophers seldom if ever ask: "What does the Libet experiment show?" Instead, they ask: "What does the experiment show *about free will*?" The first question is open-

ended and demands that we evaluate RLE on their own terms. The second question demands that we evaluate the quite perspicuous experimental data generated by Libet in terms of the vagaries of the concept of free will. We will find that little is gained by discussing RLE in terms of their bearing on "whether we have free will": for the very question concedes too much validity to assumptions about the nature of agency, and indeed the composition of the mind, that RLE call into question. For the same reason, little is gained by evaluating RLE in terms of their bearing on "incompatibilism" or "libertarianism." Indeed, nothing is gained, and much is lost, by trying to press RLE into the service of a pre-existing philosophical agenda. Most importantly, nothing is gained by responding defensively to RLE--- by trying to diminish or circumscribe their significance. RLE clearly *are* significant, and they almost certainly do generalize, granting that it is an open question how they do so.

The plan of the present paper is as follows. In Section I, we will briefly discuss LE and RLE. Here we will pose a few questions, to be decided on the basis of the material put forth later sections, and also make some clarificatory remarks. In Section II, we will discuss some of the studies that were conducted in response to LE. Many such studies have been conducted; we will focus on the most important ones. In Section III, we will discuss philosophers' responses to RLE. Finally, in Section IV, we will put forth our own interpretation of RLE. The model we will put forth is both highly schematic and highly conjectural; but if nothing else, it represents a good faith to model the quite striking data generated by LE.

Section I: Libet's Experiment: A Brief Discussion

Question 1: Is W the moment when the subject decides to move his finger or is it the moment when he becomes aware of an already-made decision to do so?

The likely answer: It is when he becomes aware of his decision; it is not when he makes that decision.

Explanation: RP is obviously the proximal psychological cause of the action. An intention is the proximal psychological cause of the action. Therefore, RP is the intention; therefore, W is not the intention; it is an awareness of the intention.

Also, by the time W happens, RP is almost complete. So even if W is a *part* of the intention that led to the action, it is a subordinate part: at most, it has a facilitating function. But intentions—and therefore decisions (acts of intention-creation)—are indeed supposed to have a *decisive* role in the genesis of the corresponding actions, and W clearly does not have such a role. If we simply insisted that W was the relevant intention, we would be taking the incoherent view that intentions did not have a decisive role in the genesis of the corresponding actions.

This is not to say that W is a mere "epiphenomenon": it demonstrably isn't.[1] For even if it is a subordinate actor, or a non-actor, in this particular context, it clearly does have effects. Consciousness of one's intentions necessarily has effects: If I am aware of a given intention of mine, my subsequent thought and behavior will necessarily be different from how they would otherwise have been. Indeed, W, though not the intention responsible for the finger-movement in question, clearly *does* have a hand in generating other intentions down the road. (This is trivially true of any intentions that he subsequently to learn about W itself, and it almost certainly holds of intentions that do not fall into this category.) Be all of this as it may, W is not the cause of the

[1] As we will discuss later, Wegner (2004) takes the position they are indeed mere epiphenomena.

finger-movement and it is therefore not the operative intention. W presupposes the existence of that intention and is therefore either an awareness of it or a decidedly subordinate part of it.

And the evidence is that W is *not* a *part* of that intention. First of all, W happens very late in the game: whatever it is that is responsible for the finger-movement is well under way by the time W makes an appearance. Mele (2009) takes the position that, for all we know, the pre-W part of the RP is the intention-buildup, not the intention proper, W being the being the intention proper. But if this were any other context, such a position would be absurd. Suppose the graph in Figure 1 concerned the growth in value of Warren Buffet's stock portfolio and that W were some unspecified but related event, e.g., Buffet's buying shares of Microsoft. If we said that *W* was what put Buffet in in the black, our position would have no plausibility at all. In fact, we would have no reason to regard W as doing *anything* to contribute to the growth in Buffet's wealth, other than, at most, simply helping to continue it.

And even this position has little support. In both Figures 1 and 2, W either has no effect, or a slight negative effect, on the trajectory of the slope. In Figure 2, it has no effect: it is not an inflection-point of any kind. In Figure 1, it is a slight *negative* inflection point: right after it, the graph's slope decreases slightly. If W were indeed the decision responsible for the finger-movement, the graph would inflect upwards after W: it would not fail to inflect; nor it would inflect negatively. Judging by Figure 1, W's role is, if anything, is to curb the causally effective process. Judging by Figure 2, W's role is indeterminate but clearly not decisive. Either way, W cannot possibly be an intention, let alone a decision, since such things are of their nature decisive causes.

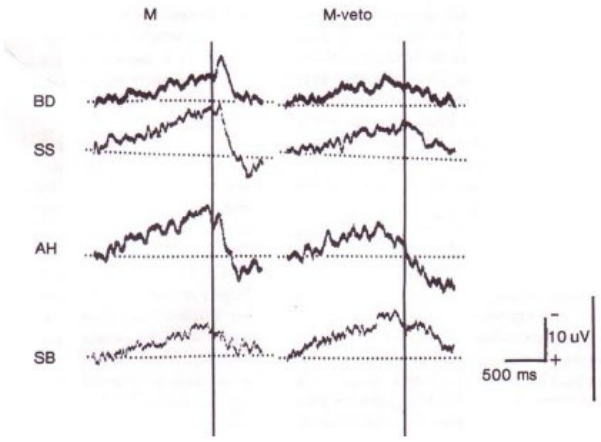
Question 2: What causal role does W have? More simply, what does W do? One answer is: W is the operative intention. But we've seen that this is false.

Another answer is: W does nothing; it is an "epiphenomenon." This is the answer put forth by Wegner (2004) on the basis of RLE. But this too is false, as we've seen; for even if W is epiphenomenal with respect to that particular finger-movement, it is still operative in other contexts.

Libet's Veto Experiment: Libet himself put forth a third answer. Like us, Libet believed RLE to show that that W is not the operative intention and, moreover, that the operative intention is the unconscious activity corresponding to the pre-W part of the curve (Libet et al. 1983; Libet 1993). This being so, Libet wanted to determine whether W at least had the power to "veto" the action in question—to put the breaks on it, as it were.

To this end, he devised an experiment (henceforth "LE2') in which the subject is given two instructions (Libet 1999; Batthyany 2009). First, he has to form an intention to push a button at a given time (e.g., when the dial on the clock in front of him is at 30); second, just before the time in question, this being when he feels the urge to push the button, he has to *refrain* from doing so. The results of LE2 are shown in in the "M-veto" (movement-veto) column of Figure 3.

Figure 3



Taken from Batthyany (2009).

The M-veto graph is much less steep than the M graph but otherwise similar. The difference in steepness presumably corresponds to the fact that, in the M-veto case, the subject's intention to push the button was tempered by a contradictory mandate. It may also correspond to the fact that, in the case of the planned action, the RP happens over a longer period of time.

Libet took the results of LE2 ("RLE2") to indicate that W has the power to veto the action in question (Libet 1999). This certainly could be the right interpretation, judging by the relative positions, in the M-Veto graphs, of W and the ensuing action. Also, in LE2, the subject was explicitly ordered to refrain from acting *in response to* W, with the result that W *necessarily* had a hand in vetoing the finger-movement. Consequently, RLE2 provides only weak support for the position that *in general* W has a veto function.

In any case, RLE2 are ambiguous, and we don't have the time to go through them in detail, but they are certainly consistent with, though not probative of, the supposition that W "polices" the subject's intention to move and that, in general, our "conscious intentions"—in other words, our conscious *awareness* of our intentions---have the power to check and possibly otherwise modify the execution those intentions. For the reasons discussed earlier, such awarenesses probably also discharge a number of other functions, and therefore not "free wheels", granting that they are not identical with the corresponding intentions.

Question 3: Do we have "conscious free will"? According to Libet, RLE show that "we do not have conscious free will" (Libet et al. 1983). By this, Libet seems to mean that our decisions are not constituents of consciousness. If this is indeed what Libet means, then he's right, at least with respect to the class of actions represented by LE. For RLE do indeed show, if

only with respect to that particular action-class, that our decisions are not constituents of consciousness. This in turn suggests that the conscious events that we *believe* to be decisions are mere awarenesses thereof.

How much does this result generalize? This is not a question that can be answered off-the-cuff, but we can volunteer a few potentially helpful remarks. First of all, the intentions with which RLE are directly concerned involve real-time engagements with one's immediate environment. Such intentions govern actions that are motivated the contents of one's current sense-perceptions and whose outcomes will have immediate and well-defined consequences for one's perceptions. Not all intentions are like this: my intention to write a book over the next three years does not fall into this category; nor does my intention to start a company; nor indeed does my intention to have a Zoom meeting later today. In general, actions driven by so-called "standing intentions"---intentions that do not mediate real-time engagements with one's environment—fall outside the action-class with which RLE are directly concerned.

Consciousness is presumably *relatively* heavily involved in the guidance of real-time engagements with one's environment. And given that intentions are not constituents of consciousness where real-time actions are concerned, it is doubly unlikely that they are constituents of actions where other forms of action are concerned. Consider my intention to build a digital trading platform for sugar futures brokers. This intention was not the result of the contents of my sense-perceptions in some specific situation; it developed on the basis of judgments made in response to a wide range of situations and linked together by protracted and therefore largely unconscious thought-processes. It would therefore be extraordinarily odd if intentions of *this* kind were identical with constituents of consciousness, given that the intentions governing real-time actions are not such constituents.

Indeed, a standing intention demonstrably cannot possibly be a constituent of consciousness. My intention to create a new trading platform is an enduring structure, like my knowledge that 2+2=4. I can be conscious of it, but it is not itself a constituent of consciousness. If *any* intentions are constituents of consciousness, it is those that govern perception-mediated real-time engagements with one's environment and it is therefore those with which LE is directly concerned.

So yes, Libet's experiment *does* show that we do not have "conscious free will"; for it shows that our decisions, even those that govern real-time responses to the environment, are not constituents of consciousness and, moreover, that we are not even conscious *of* them until after an appreciable time-lapse (.2 seconds).

Question: Is this time-lapse garden-variety 'tape-delay'? In other words, is the time-delay as small as it could be, given that the intention precedes the awareness of it or, on the contrary, is this time-a reflection of W's function? This is an empirical question and I don't know the answer, but my suspicion is that it is not mere tape-delay. Consciousness operates extremely quickly. The average sprinter takes around five steps/second—.2 seconds/step--and he is, or at least can be, conscious of each step. In fact, he is, or at can be, conscious of step-components, each of which takes a mere fraction of that .2 seconds. This being so, the .2 second delay between intention and awareness is very possibly *not* mere tape-delay. If so, this reinforces the idea that W serves less as an action-instigator and more as an action-monitor.

Section II: Follow-up Studies

Some Representative Studies

In 1998, a team of researchers, headed by Patrick Haagard, conducted a variant of Libet's experiment (Haggard& Eimer 1999; Haggard 2011). There were two parts to Haagard's experiment. The first involved simply replicating Libet's experiment. The other was a variant of Libet's experiment in which the subject could choose to push the button with either his right index finger or his left index finger. (In Libet's experiment, it will be recalled, the subject could only use one index finger.) Whereas the first part measured readiness potentials, the second measured so-called "lateralized" readiness-potentials (LRP's), these being readiness-potentials corresponding to actions that involve decisions as to which side of one's body to move. The first part of Haagard's experiment confirmed Libet's findings. The second part showed that in the relevant respects LRP's are virtually indistinguishable from RP's: they begin at the same time, relative to both W and the finger-movement, and they have the same trajectories (Haggard& Eimer 1999; Haggard 2011).

In 2008, a team of researchers, headed by Chun Siong Soon, conducted a study similar to Libet's, except that their study used much more sophisticated brain-scanning technology (Soon et al. 2008). Libet's study used an EEG (electroencephalogram), which measures electrical activity in the brain on the basis of temperature-changes in the subject's scalp. Soon's study used functional magnetic resonance imaging (fMRI), which measures brain activity on the basis of changes in oxygenation levels in the brain. fMRI is much more accurate and high-resolution than EEG. Soon's team found that the relevant RP's—the ones that led to the subject's finger-movement and therefore constituted his intention to do so---began much earlier than Libet's experiment had indicated, sometimes preceding W by as much as 10 seconds (Soon et al. 2008; Fried et al. 2011).

In 2010, a team of researchers, headed by Itzhak Fried, conducted a variant of Libet's study in which depth-electrodes, as opposed to EEG or fMRI, were used to determine the onset-time of the RP's in question (Fried et al. 2011). This made it possible to identify the very first neural firing involved in the RP in question and to identify the "neuronal recruitment" process whereby additional neural firings came to be involved. Fried found that the relevant RP began 1500 ms (1.5 seconds) prior to W---1.3 seconds prior to the time at which Libet, operating with less sensitive technology, detected RP-onset.

Finally, in 2012, a team of researchers, led by Aaron Schurger, conducted a different variant of Libet's study (Schurger et al. 2012). In Schurger's study, the subject would hear randomly distributed clicking noises during the interval between RP-onset and W. These clicks were reasonably loud and were intended to be—and presumably were—distracting. Schurger's intention was to determine what influence, if any, these clicking sounds had on the relevant variables, these being the shape of the RP and the positions, relative to each other to RP-onset, of W and the finger-movement. Schurger found that none of these were appreciably changed. In other words, Schurger's experiment generated much the same results as Libet's study, the distracting clicks notwithstanding. This suggests—and Schurger takes it to suggest—that, prior to W, the RP in question is robust and is overwhelmingly likely to eventuate in action (Schurger et al. 2012).

Discussion

Other follow-up studies have been conducted; none have impugned RLE and most if not all have confirmed RLE (Pollen 2006). Indeed, as we have seen, these studies suggest that Libet was even *more* right than he thought. The operative RP is initiated even earlier than Libet

thought, and, once initiated, it is overwhelmingly likely to result in action (Fried et al. 2011; Braun et al. 2022).

Considered on their own terms, these results are quite straightforward. What we believe to be choices that are initiated in consciousness are not initiated in consciousness; they are initiated prior to any conscious awareness of them and also far outside any such awareness.

To be sure, it may be unclear how these results bear on the question "do we have free will?" But that would seem to be more a reflection of that question than it is of those results; for it isn't at all clear what it means to say that someone has "free will." Also, it is assumed that, if there is such a thing as free will, it is "selves" that have it, and it isn't clear what a self is. Nor is it clear whether there is a self wherever there is sentience or, on the contrary, whether sentience must configured in a certain way for there to be a self. So yes, it is unclear how these experimental findings bear on questions relating to free will and selfhood, but only because those concepts are unclear, not because those findings are unclear.

In the next section, we will discuss how philosophers have responded to these findings. By way of anticipation, we will find that philosophers have responded very conservatively and defensively to these findings, with the result that they have done more to obfuscate than to illuminate their significance.

Section III: Philosophers Respond to Libet

Preliminaries

Determinism: The doctrine that the state of the universe at any given time fixes the state of the universe at any later time.

Indeterminism: The doctrine that determinism is false.

Compatibilism: The doctrine that free will is compatible with determinism.

Incompatibilism (also known as "libertarianism"): the doctrine that compatibilism is false.

The motivation for incompatibilism: If determinism is the case, then what any given person is doing on any given occasion the inevitable consequence of the state of the universe before that person's birth, and it is hard to someone could be free under such circumstances.

The motivation for compatibilism: First of all, to the extent that the *external* world is indeterministic, one's ability to control it, and therefore one's freedom (if one has any), is abridged. If I dial a number on my phone and there is a 72% that the call goes through, then I have less control over the world, other factors equal, than I do if there is an 84% chance that it goes through.

Similarly, one lacks control over oneself, and one therefore lacks freedom (if one has any), to the extent that the *internal* world is indeterministic, i.e., to the extent that there is indeterminism in the event-system my mind-brain and nervous system. If there is only 72% chance that my fingers move when I want them to move, I have less control over myself, and therefore over the world, than I do if, other factors equal, there is an 84% chance that they move when I want them to. So determinism is necessary to have power over oneself and therefore and therefore to have freedom, whatever that is exactly.

"What about our thought-processes?", it will be asked. "If they are predetermined, then surely they aren't free." But what about if they are *not* predetermined? In that case, they come out of nowhere; they aren't rooted in one's beliefs, values, or intentions; they are mere visitations. In fact, if thoughts so totally lacked a causal basis in a given person's pre-existing

psychological structures, along with the physiological structures mediating them, then it is unclear why those thoughts would be *his* thoughts; they would, it would seem, simply be free-floating thoughts. In any case, thoughts that lacked a causal basis in one's pre-existing condition would seem in the nature of intrusions—similar to an obsessive-compulsive's "intrusive thoughts." Though uncaused, they would not from that person exercising his agency; indeed, they would be *undergone* by him and therefore would neither give him freedom nor represent deployments of any freedom that he might have.

So, we only have freedom to the extent that the world, including ourselves, are deterministic.

Conclusion: Indeterminism threatens freedom; indeed, there doesn't seem to any respect in which any form of indeterminism makes us free. It may be that determinism also threatens freedom; but determinism seems to be a prerequisite for freedom, at least in so far as it is possible to have freedom; moreover, there does seem to be at least *a* kind of freedom that is compatible with determinism. I am acting freely, at least after a fashion, when my actions align with my values and, more generally, when they represent a fulfillment of my biological program. If, in making a decision, I am exercising my rationality to the greatest extent possible and---what may or may not be included in that—I am also choosing a path that will enhance me, rather than diminish me, then, at least arguably, I am in that context acting freely. In any case, this is the closest thing to freedom that is possible or even conceivable.

All of this being so, if it isn't clear how results of Libet's study (RLE) bear on questions about free will, that is because the concept of free will is unclear. RLE are quite clear: When it comes to a certain subclass of actions involved in real-time engagements with the environment, intention-mediating RP's predate the conscious events that people are inclined to believe their

intentions to be. It may not be exactly what we ought to infer about the mind from it, but there is nothing unclear about this fact *per se*.

Philosophers' Responses to the Libet Experiment

We will consider a few works written by philosophers concerning the Libet Experiment. They are highly representative of the philosophical literature in general.[2]

"How does the neuroscience of decision-making bear on our understanding of moral responsibility and free will" (2010) by Adina Roskies.

In this paper, the author doesn't address the Libet experiment explicitly. (At the same time, the bibliography contains an article of hers on the Libet experiment). Rather, her concern is the philosophical significance of neuroscience as a whole. Her position is that it has none. Neuroscience, Rouskies says, merely shows us how intentions, beliefs, and so on are physiologically *implemented*, but it doesn't otherwise give us information about them, and it certainly doesn't give us any information that warrant a revision of our pretheoretic judgments about them. Indeed, Rouskies go so far as to say that when neuroscientific results bear on philosophical questions concerning, they "dovetail nicely" (Rouskies 2012, p. 1025). with traditional philosophical views, specifically with traditional philosophical views concerning agency, the primary one being that it must be understood in compatibilist terms.

Although Rouskies doesn't mention Libet by name, she obliquely addresses his work:

[2] Unlike the philosophers that we are about to discuss, Dennett (2015) takes RLE at face value. But he is exceptional in this respect.

> Neural representations of aspects of the world and of internal variables constitute information that is integrated over time by decision mechanisms. In general, decisions occur when firing rates of neurons representing one of the choice options reach a certain threshold. This seems to involve fairly low-level processes. But human choice and decision-making seem to involve more active processing and modulation by the agent. How do high-level control processes such as attention, valuation, policy commitments, and so on affect the basic machinery of decision-making? That is a wide open question, begging for concerted research.[3]

The idea here seems to be that the Libet Experiment is concerned only with "low level", semi-automatic actions; it has nothing to do with non-automatic actions, such as my decision to start a hedge fund or build an addition to my house. According to Rouskies, these actions, for lack of a better term, are almost in the same category as mere reflexes; elsewhere in the article, Rouskies describes them as "degenerate actions." The implication is that we don't really learn much of anything about *bona fide* intentions by studying them. The further implication is that this is but one more illustration of how little neuroscience has to teach us about the structure of the mind.

But the actions with which RLE are concerned are not reflexes: unlike reflexes, they are driven by intentions. And RLE *do* give us new and counterintuitive information about those intentions. What RLE show is that the conscious events that we thought to be the very essence of

[3] Rouskies 2012, p. 1025.

decision-making are either not involved, peripherally involved, or involved totally differently from how we previously thought.

For this very reason, the idea that neuroscientific research "dovetails nicely" with compatibilist views of agency is misleading. According to compatibilism, *you* are free even though *your* actions are predetermined. But Libet's experiment shows that your center of consciousness—which seems to be central to your very identity---has a rather oblique and tenuous relationship to paradigm cases of "free" action. Does LE definitively show that we don't have "free will", whatever that means? No—but it does definitively show that one's center of consciousness—this being what people intuitively regard as the center of their very selves---has a much more subordinate role in decision-making that people, including philosophers, assume.

Rouskies asks: "How do high-level control processes such as attention, valuation, policy commitments, and so on affect the basic machinery of decision-making?" Her answer: "That is a wide open question, begging for concerted research." Yes, it is an open question, but contrary to what Rouskies is implying, there is no reason to believe it will do a better job than RLE of validating compatibilism or any other pre-existing philosophical doctrine concerning agency.

"A Critique of Libet and Wegner's Argument Against Free Will" (2023) by Ferdinard Fosu-Blankson.

According to the author, Libet claims that his experiment proves there to be no such thing as "free will", and the author's purpose in this paper is to show that Libet's experiment fails to prove that.

There are two points to make right away. First, Libet claims that his experiment proves there to be no such thing as "conscious free will", by which Libet appears to mean that, although

we may have free will, the conscious events that we believe to be our decisions are, at last in some cases, not our in fact our decisions.

Second, and more importantly, Libet's primary purpose is to identify a fact about the mind, not evaluate that fact philosophically. Maybe RLE show that we don't have free will; maybe they are irrelevant. Either way, RLE identify a fact about a certain class of decisions; and that fact is non-trivial, granting that its significance is unclear.

In any case, in his attempt to refute the position that he ascribes to Libet, the author makes the following claims.

Claim 1: According to Fosu-Blankson, Libet is guilty of over-generalization (or "hasty generalization", to use his term).[4] As we've noted, there is a respect in which RLE are circumscribed; they only directly apply real-time to one's environment. But this isn't what Fosu-Blankson means; his position is that RLE only apply to finger-movements and therefore cannot be generalized to other body-movements.

Analysis: This is an extraordinarily weak point. It would be profoundly odd RLE held with respect to finger-movements but not to leg-movements are arm-movements. If Libet is overgeneralizing, it isn't because he believes his results to hold with respect to other body-parts; it is because not all intentions mediate real-time engagements with one's environment and Libet is assuming otherwise. Also, as we've discussed, Libet is ultimately not guilty of overgeneralizing. LE is set up in such a way as to maximize the role of consciousness in the actions in question. The subject is wide awake; he is given explicit instructions; he can push the button whenever he wants within the relevant multi-minute time-window. In LE, the subject is not told to write a novel or come up with a business plan. He is not told to do anything that

[4] Fosu-Blankson 2023, p. 3.

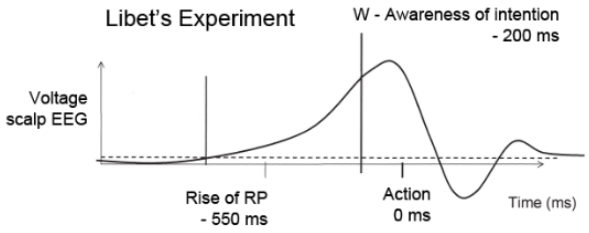
would involve protracted thought or that would therefore rely disproportionately on unconscious thought. What he is told to do is about as consciousness-centered as an action can be.

Claim 2: Libet is confusing correlation with causation. Yes—the RP precedes the finger-movement and it also precedes W, but that doesn't mean that it causes either.[5]

Analysis: The RP obviously does cause the finger-movement; either that or it is temporally coincident with whatever the cause is. Either way, W is not the operative intention, this being the relevant fact.

As we will see, the position that Libet is confusing correlation with causation—that we don't really know that the RP causes the finger-movement—is one that occurs repeatedly in the philosophical literature in LE, even though there is no evidence for it and strong evidence to the contrary.

Let us once again look at Figure 1:



The finger-movement happens right at the top of the curve, at which point it plummets. Clearly, the curve tracks whatever it is that causes the finger-movement. And whatever this cause is, W is not an important part of it or simply not a part of it.

Claim 3: Libet's experimental data is questionable. To quote Fosu-Blankson:

[5] Fosu-Blankson 2023, p. 12.

In Libet's and Wegner's studies, they engage in empirical (scientific) processes. Libet tests the presence of RPs, their timing, and efficacy during voluntary acts, while Wegner formulates the theory of apparent mental causation in an attempt to establish the unconscious system responsible for human actions rather than (conscious) human agency. However, these experimentations cannot be said to yield accurate picture of what transpired in the conscious configuration of the participants of these empirical studies.[6]

Libet's results have been duplicated many times over (Pollen 2004, 2006; Braun et al. 2021). Yes, there can be margin-of-error issues in experiments, but there is no reason to believe such problems are operative in this context.

"Free Will and Neuroscience: Decision Times and the Point of No Return" (2019) by Alfred Mele.

Mele starts by attributing the following argument to Libet[7]:

1. In Libet-style experiments, all the decisions to act on which data are gathered are made unconsciously.
2. So probably all decisions to act are made unconsciously.
3. A decision is freely made only if it is consciously made.
4. So probably no decisions to act are freely made.

[6] Fosu-Blankson 2023, p. 4.
[7] Directly quoted from Mele 2019, p. 84.

Libet's objective is to prove that this argument fails and, more generally, that Libet's experiment prove nothing of philosophical significance.

First a preliminary point. Mele is making it sound as though Libet was writing a philosophy paper and that the experimental data he generated is simply an attempt to produce an empirical proof of a philosophical thesis. But that is not the case; and even if it were, Libet's results would still be worth understanding on their own terms.

In any case, Mele provides three arguments for his position.

Argument 1: Libet's experimental data cannot be trusted. We don't know exactly when W occurs; we only know when the subject says that W occurs. When we take into account the gap between these two events, and also between W and the subject's resolution reporting having W, we don't find there to be any reason to deny that W is coeval with the operative intention.[8]

Analysis: This is certainly a *possible* criticism of Libet's experiment, but there is little reason to believe that it is anything more than this. The subject established the time of W by coordinating it with the position of a dial on a clock that made one full rotation every 500 ms (.5 seconds). In order for RLE to fall into margin-of-error territory, the subject would have to be very wrong about when W occurs; he would have to have report having W when the dial was at 0 when, in actuality, the dial was at 40. Although it's theoretically possible that this is what happened, there is no reason to believe in the absence of independent corroborating evidence, of which Mele provides none. Also, Libet's findings are confirmed by later studies in which more sensitive technology was used; and in such studies, the gap between RP and W is simply too large to chalk up to error-margin issues.

---

[8] Mele 2019, p. 86

Argument 2: RLE only concern "proximal intentions" (intentions to perform actions in furtherance of real-time engagements with the environment, e.g., my intention to run away from the growling dog that is running towards me) not "distal intentions" (intentions to perform actions in the future or over a period of time, e.g., my intention to write a dissertation).[9]

Analysis: The distinction between proximal and distal intentions is completely valid. Moreover, Mele is right RLE only directly concern proximal intentions. But for reasons that we have already discussed, distal intentions are extremely unlikely to be *more* consciousness-based than proximal intentions. In fact, because distal intentions are of their nature relatively enduring psychological structures, as opposed to mental events, they cannot possibly be identical with the ephemera that populate consciousness. Theoretically, it is possible that the decisions that lead to such intentions are categorically identical with conscious events, but there is no evidence to suggest this.

In fact, there is strong evidence to the contrary. I often find myself carrying out intentions that I did not consciously form. Was there necessarily a "moment" when I decided to study finance or when I decided to invent a new trading platform? Sometimes distal intentions result, or at least seem to result, from conscious epiphanies, but they do not appear to do so categorically.

Argument 3: The 'actions' studied by Libet are so low-level that they are more in the nature of reflexes than *bona fide* actions; and little if anything is to be inferred from them about the nature of intentions. Rouskies (2010) made this very point, it will be recalled, and it shows up repeatedly in the philosophical literature.

---

[9] Mele 2019, p. 87.

Analysis: The behaviors with which RLE are concerned are textbook cases of *bona fide* actions---of non-automatic, non-reflexive, "feely made" (whatever that means) actions. They are simple, yes, but not reflexive otherwise unintentional. Moreover, those actions are textbook cases of what we would ordinarily regard as "conscious." The subject was free to choose when to push the button; he was not given incentives to push it at this as opposed to that time; the context was not emotionally driven; the subject was not given subliminal messages; he was not manipulated, conditioned, or programmed, and his consciousness was not otherwise circumvented. The experiment was set up to maximize the role of consciousness in the relevant action. Consequently, the behaviors in question were the exact opposite of what Mele is alleging them be.

"Evidentiating Free Will" (2014) by Nicholas Rescher.

Rescher puts forth two arguments to the effect that RLE prove nothing of substance.

Argument 1: The suite events that Libet is studying---the event-chain comprising RP and W---does not consist of "parts", and we therefore cannot say that some early "part" of it is responsible for some later "part" of it:

> RPI [reaction potential increase, i.e., the activity represented by the upward slope of the graph] is simply the first part of a single, indivisible unit of a process that is temporally indissoluble but only the last part of which rises above the level of conscious awareness. The transaction involved in these minuscule temporal processes is *atomic* in the sense of the ancient Greeks: it is indivisible. One can *distinguish* among parts of it, but one cannot *separate* them. For such an integrated whole is not separable into two distinct components constituted by an inclination-to-act separable from the act itself but, rather, one unit so compressed

> in time that its stages, while conceptually distinguishable, are not processually separate.[10]

In other words, because RPI is "indissoluble", the conscious event at W is no less causative of the resulting finger-movement than any of the previous parts of RPI.

Analysis: This is question-begging. RLE show that W is *not* as causally important as it was previously thought to be, and it therefore shows that philosophical views, such as Rescher's, are empirically false. Also, most causality is "processual", meaning that it is mediated by processes. A light-beam is a text-book case of a causal process; so are a gust of wind, a rolling ball, and a growing fetus. The RP is a textbook case of such a process, and it is absurd to suppose that the events composing the RP at a given time are not the effects of the events composing the RP at earlier times.

Argument 2: According to Rescher, Libet is confusing correlation with causation. To quote Rescher (2014):

> Where occurrences are not punctiformly instantaneous in time but come in small transtemporal droplets, this coordinative situation can also obtain with the earlier part of the unity exerting no *causal* priority over the latter. One must distinguish between backward causation (a scientific bugaboo) and noncausal connectivity (to which science neither does nor can have valid objection). (If today is Monday, yesterday has to be Tuesday, but the one does not *cause* the other either way.)[11]

[10] Rescher 2014, p. 96.
[11] Rescher 2014, p. 96.

The earlier parts of the RP clearly do "exert causal priority" over the later phases, and "causal connectivity" clearly does hold between them. The comparison with days of the week is misconceived. The references to "backward causation" are red herrings.

"Does the brain 'initiate' freely willed processes? A philosophy of science critique of Libet-type experiments and their interpretation" (2013) by Hans Radder and Gerben Meynen.

The authors argue the RP does not cause W and also does not cause the subsequent finger-movement. They put forth two arguments for this contention.

Argument 1: RP-onset is not an INUS-condition for the occurrence, .2 seconds later, of W. (x is an INUS condition for y iff x is an insufficient but non-redundant part of an unnecessary but sufficient condition for y.) The reason: even if RP is sufficient for the occurrences, at *some later time*, of W, it is not sufficient—or at least not there is no data to indicate that it is sufficient---for the occurrence of W *.2 seconds later.* For similar reasons, the RP does not cause the finger-movement.

To quote the authors:

> [T]he experiment has produced no data at all to support the claim that the occurrence of a specific set of events *RP-t1* ∧ *Xbr* is sufficient for inferring the specific *timing t2* of *W-t2*. Put differently, the experimental results do not explain why, in a particular trial, the conscious free will occurred at specific time *t2*, rather than somewhat earlier or later (see also n. 12, and the corresponding discussion in the main text, below). Furthermore, since the possible role of *Xbr* has not been concretely investigated, Libet's experimental design has not

provided any evidence for or against the claim that *Xbr* alone is sufficient for *W-t2*. For this reason, we do not know whether *RP-t1* is non-redundant.[12]

Analysis: There is a wealth of experimental data to support the exact claim that the authors say is not supported. When the Libet experiment perfectly replicated, the delay is consistently in the neighborhood of .2 seconds. When the experiment in some respect, e.g. the subject can click with either his right or left index finger, the gap between RP-onset and W, though not necessarily the same as in the original version of LE, is consistent.[13] Ultimately, the authors are simply saying that, because the relative times of the events in question are not measured with quantum accuracy, we cannot establish causality—which is absurd.

Argument 2: In order to establish that x causes y, we need to establish that changes in x-like events lead to changes in y-like events; we need to know that input-changes are reflected in output-changes. We cannot possibly know this where RP and W are concerned, and we therefore cannot assert that RP-onset causes W. To quote the authors:

> In the Libet experiment, the claim that *RP-t1* is the cause of *W-t2* requires that specific changes in the timing of the conscious free will to act can be brought about by certain interventions that change the readiness potential. In the manipulability theory, such interventions are not limited to changes brought about by human experimenters (Woodward, 2003a, p. 94). In principle, their occurrence may also be the result of some natural process, what Woodward calls a "natural

[12] Radder & Meynen, p. 9.

[13] Braun et al. 2021.

> experiment." In Libet's experiment, however, such interventions (be they human or natural) have not taken place. Moreover, according to some authors, the relevant human interventions (which would effect a change of the timing of W by modifying RP) cannot be performed. As Haggard concludes, "the problem of describing and classifying the brain processes of action initiation" is that "we cannot intervene on the processes of action initiation and see how RPs and conscious experience are altered in consequence" (Haggard & Libet, 2001, p. 62). Such an intervention or manipulation, however, is exactly what is called for in order to satisfy Woodward's manipulation requirement. Furthermore, in the Libet experiment other factors that might account for changes in the timings of W have not been considered either. Hence, we have to conclude once more that Libet's experimental results cannot be used to validate the claim that *RP-t1* is the cause of *W-t2*.[14]

Analysis: First of all, this argument at most establishes that we cannot know that RP-onset causes W. Second, for the reason previously given, we *do* know that changes in RP are reflected in changes in W, including when W happens relative to RP. Many different Libet-style experiments have been conducted; and they've shown that the length of the RP, as well as the relative positions of RP-onset and W, depend in a consistent manner on the exact task that subject is being asked to perform.[15]

Discussion: This article is not significant; the arguments that it puts forth, when not prohibitively vague, are demonstrably false. But the article provides an excellent illustration of

[14] Radder & Meynen, p. 10.
[15] Radder & Meynen, p. 10.

the hyper-defensive manner in which philosophers have tended to response to the Libet experiment.

Some other recurring philosophical tropes

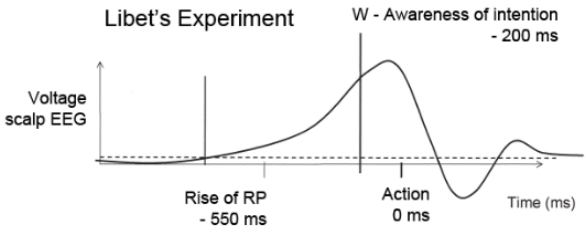
There are some recurring themes in the philosophical literature concerning the Libet Experiment that we have not yet mentioned or only alluded.

Libet confuses urges with intentions: According to this line of thought, W *is* indeed the operative decision. Yes—W results from the RP. But prior to W, RP is a mere urge and therefore not anything that will by itself lead to action. W is needed to convert this this intention-precursor into an actual intention, and W is therefore precisely what we always thought it was. To quote Knowles:

> Libet asks participants to indicate when they feel the desire to perform an act and discovers that an electric activity in certain brain areas precedes the movements participants make. He refers to it as readiness potential, RP, since it is registered before the activation of participants' muscles representing the desired movement. Libet notes that RP was registered well in advance of movement and well before subjects become "aware of the wish or urge to act", which approximates at about 550 msec. prior to an act. Libet interprets these results to mean that "the initiation of the voluntary act appears to begin in the brain unconsciously, well before the person consciously knows that he wants toact!". (Libet: 1999, p. 51) Well, Libet

misinterprets and misidentifies the participants' brain processes as readiness to act rather than as desire to act.[16]

Analysis: If W is indeed a decision, then it has a decisive role in causing the finger-movement. But it doesn't. Again, let us look at Figure 1:



W does not have a decisive role in causing the finger-movement. The finger-movement happens at the vertex of the curve, and it happens because that curve reached its vertex. Whatever W's role in this is, it is not decisive, and it therefore isn't a decision. In fact, judging by the graph, W is opposite of decisive; it is in the last place that a decision to act would be. If W happened at the very beginning of the slope, it could be a decision; for then it would clearly be what caused the finger-movement. If it happened at the peak of the slope, it could be a decision; for then it might be what converted an erstwhile urge into a veritable intention. But W happens far too late to launch the process and too early to transform it.

Suppose that a statistician or physicist or biologist were looking at this graph but he didn't what it was a graph of: he only knew that it represented *some* process or other. Relative to this information, the probability he would assign to W's being transformative or otherwise

[16] Knowles 2022, p. 12.

decisive would be minimal. So, to say that W *is* transformative simply contravenes the data, and to say that it *must* be transformative simply begs the question.

Libet's results don't generalize because they concern actions that fail to engage the subject's rationality: According to Bayne (2006), LE is concerned with actions that require no thought or deliberation and therefore only minimally engage the subject's agency. (The subject merely has to push a button when he feels like pushing a button; he doesn't have to think or plan.) For this reason, Bayne claims, RLE tell us little or nothing about human agency:

> [A]lthough Libet-actions involve an act of will they do not involve deliberation – at least, not immediately prior to the action. In my terms, they are examples of disinterested agency, for the agent has no reason to flex their wrist at one particular time rather than the other, or to flex it in one way rather than another. Indeed, Libet-experiments are explicitly constructed so as to minimize the rational constraints under which the subject acts. We might think of Libet-actions as manifesting the liberty of indifference.[17]

Immediately thereafter, Bayne writes:

> With the foregoing in hand, let us return to the question of whether Libet-actions are paradigms of free will (as we intuitively conceive of it). Are disinterested actions our central exemplars of free will, or does that epithet belong to deliberative actions? Philosophers do not agree on the answer to this question, and

[17] Bayne 2011, p. 30.

> the systemic research that would be required in order to settle this dispute has not been carried out. That being said, my suspicion is that Roskies is right to identify the central or core cases of free will – at least, the kind of free will that is most intimately related to moral agency – with deliberation and rational reflection.[18]

In a word, RLE is concerned with actions that fail to engage rationality or agency. Therefore, we cannot deduce anything from RLE about human action in general.

Analysis: First of all, there is no experimental data to support Bayne's claim. Second, for reasons already discussed, the more deliberation and "rational reflection" are involved in an action, the less likely the operative intention is to be fully conscious. Also, the actions studied the Libet experiment are probably not as one-dimensional as Bayne and Rouskies make them out to be. If you ask me to push a button at random intervals, I have to use more thought to follow your instructions than I would if you asked me to push a button at regular intervals. In the second case, I can let myself go on autopilot; the first case, I cannot.

It is a "phenomenological datum" that we have free will: According to this position, we "just know" that we have free will; indeed, such knowledge is etched into the very structure of the human mind. Foso-Blankson writes:

> Wegner deploys a theory based on experimentation that studied human actions and drew conclusions that infer the conscious will's incapability of causing any effect in the brain or any part of the body such that it will result in action. The realist position is that ontology is not contingent on conscious awareness. The fact

[18] Bayne 2011, p. 30.

> remains then that without human experience in aetiology and decision-making, there would not emerge a conception of free will. The reality of free will is experienced and not covert to our perceptual perusal.[19]

Josef Siefert writes:

> A first response taken from the immediate evidence of free will in the cogito imposes itself on us: The power of the free human person over his acts ("for if we will, it is; if we will not, it is not") is given just as immediately as his own existence in the experience of consciously living his being and performing free acts.[20]

Analysis: This position is straightforwardly question-begging. Also, the comparison with the *cogito* is misconceived. The *cogito* is to the effect I couldn't think unless I existed; as such, the truth that it registers, in addition to being thoroughly trivial, is analytic, i.e. its negation is incoherent. But there is nothing trivial about the assertion that I have free will. For free will, if it exists, is a form of causation; and it is always an open empirical question what causes what.

Section IV. A constructive response to the Libet Experiment

"If a stone became conscious while falling, it would think that it was falling of its own free will." Spinoza.

---

[19] Fosu-Blankson 2023, p. 2.
[20] Siefert 2011, p. 383.

Preliminaries

The Libet Experiment has been replicated many times, always with the same results. And similar studies, involving different measuring instruments and different tasks, have generated results that are consistent with RLE. So, while it is a theoretical possibility that RLE are spurious, it is *only* a theoretical possibility.

A more constructive approach, I would suggest, is to accept those results, if only tentatively, and try to model them. In the present section, I will try take a few tentative steps in this direction. What I say will sometimes have a speculative and even personal character.

First point: As previously stated, RLE directly establish that where some actions are concerned, the conscious events that we believe to be our intentions are not our intentions. This suggests that those conscious events are mere awarenesses of the corresponding intentions.

Next point: RLE show, quite definitively, that the conscious event that the subject experiences as being identical with his decision to move his finger is not his decision to move his finger. Indeed, it is not even involved in that decision; it is on the sidelines, at least vis-à-vis that particular decision. Thus, the consciousness on the subject's part that he identifies as the nucleus of his decision is merely an awareness of that decision, and that decision was therefore made by some other party. But of course, it is not just W that we are discussing; it is the W-analogue present in every other "conscious" action: it is the "center of consciousness" that each represents. The subject experiences this center of consciousness as being the essence of his agency, at least where "conscious" actions are concerned, but that isn't what it is.

This is a striking result, and, contrary to what some philosophers allege, it is hard to compatibilize with compatibilism. Compatibilism says that *you* are free even though *you* are

determined. Compatibilism doesn't deny that *you* act; rather, it says that how *you* act is predetermined—and that you are free so long as the operative mechanisms meet certain conditions. But what the Libet experiment shows is not so much that *you* are determined as it is that *you* are on the sidelines. By the time *you* make an appearance, the decision has already been made.

What is ominous about the Libet Experiment, that is to say, is not so much that it shows your choices to be predetermined but that it shows them not to be yours. They are yours to behold, LE shows us, but not to make.

To clarify: It is one thing if your actions resulted from conscious, albeit predetermined, decisions on your part: this is what the compatibilist believes to hold. It is a very different thing if what you thought were the operative conscious decisions are in fact non-operative: this is what the Libet Experiment shows us to hold. The Libet Experiment is not just an iteration of the very old idea that choices have causes. It is proof that we what we thought were choices aren't. It is also proof that the center of consciousness that a person regards as the nucleus of his very self, and of his agency in particular, isn't.

In a way, then, what the Libet Experiment shows us is even more radical than what Freud showed us. Freud showed us that *some* mental activity, including some decisions, are made unconsciously, sometimes with the result that there is internal conflict. But Freud was not saying that your conscious decisions, or what you believe to be your conscious decisions, are lame ducks. His position is that consciousness is like a monarch whose power, though considerable, is limited. Freud didn't deny that we take to be our conscious decisions are in fact conscious decisions; nor did he deny that that these conscious decisions represented the very nucleus of our

agency. In a word, Freud regarded this central conscious authority as, indeed, a central authority, albeit a weak one---like a ruler who, though weak, is not a puppet.[21]

But Libet's work shows that in the very contexts where this ruler least appears to be a puppet, that is precisely what he is—that when he is most certain that he is calling the shots, he is in fact a mere spectator who isn't even told what the shots are until after they have been fired. This is a dark and hard-to-absorb point, and it is not compatible with any well-developed philosophical position concerning human agency. All such positions assume that the events that we regard as our conscious choices are indeed our conscious choices, with the qualification that those choices are "free" according to some of those positions but not according to others. Libet's experiment shows this to be false, and we need to develop a conception of the mind, specifically of human agency, that accommodates this fact.

A Tentative Model

In what follows, "W" refers to an arbitrary one of the conscious events with which given person intuitively identifies the intentions governing his so-called "conscious" actions.

First of all, even if W is epiphenomenal with respect to the corresponding intention, it is not epiphenomenal *tout court.* As we've already discussed, there is some evidence that W can "veto" a pre-existing intention. But W's significance likely extends far beyond that, and we will now make a few speculative, but hopefully not totally insubstantial, remarks as to its significance.

---

[21] Freud 2005.

Consciousness is real-time awareness. There are two kinds of real-time awareness. There are real-time awarenesses of external events-- sense-perceptions, in other words. And there real-time awarenesses of internal processes—which, for the sake of brevity, we will refer to as "introspections", even though this Is not how this term is term is ordinarily used. Introspections assume many forms: they can be pains, pleasures, occurrent thoughts *about* pains and pleasures, mental images, gnawing doubts, inchoate aches, and so on. What matters in this context is that, in some cases, consciousness is inwardly directed and in others outwardly directed.

This suggests that one of the purposes of consciousness is to integrate these two bodies of information—that it is to integrate information about the external world with information about the internal world. Consciousness pools these two datasets, and decisions as to how to maximize the subject's welfare are made on that basis. Of course, the presumption is that it is consciousness that makes such decisions; but that is not the case, at least not if the results of the Libet experiment are to be trusted. What consciousness *does* do, though, is either *approve* or *disapprove* of decisions already made. Libet himself has provided some reason to believe that it can disapprove of them. I would suggest that it also actively approves of them, if only after the fact, with the qualification that these conscious acts of decision-*approval* are experiences as conscious acts of decision-*making.*

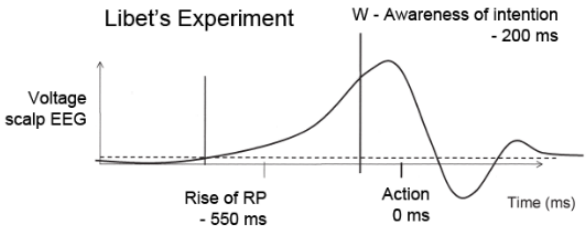
Why are these acts experienced as decision-making instead of as decision-approval? I obviously don't quite know, but one point to make is that, after a fashion, these acts *are* decisions—decisions to green-light an existing intention. Also, if were experienced as exactly what they are—if they were experienced *as* being cases of approving an intention made by some other party—that might overburden consciousness. One makes innumerably many choices within the space of a single minute; if one's conscious awarenesses of these choices were overly fine-

grained, the result would be to bog down consciousness in matters with which it needn't concern itself; and the illusion that consciousness itself the source of these decisions—the "illusion of free choice"—is a small price to pay for this gain in efficiency.

Also, to venture one last speculation, in approving of a given intention, W, though not the cause of that particular intention, does implicitly green-light future similar intentions. In other words, it tells the relevant part of one mind to carry on apace. And, by parity of reasoning, if W vetoes a given such action, it tells it to redirect its efforts. So, W does have a directive role with respect to future intentions, granting that it serves to approve or disapprove of the intention with which the subject mistakenly believes it to be identical.

Conclusion

Once again consider Figure 1:



Suppose you didn't know what this graph was a graph of: all you know, let us suppose, is that describes a growth-process of some kind. And suppose you don't know what "W" stands for; all you know about it, let us suppose, is its position on that slope. Under these circumstances, the idea that W is a crucial event in the process in question is absurd.

W is the conscious event that the subject believes to be his decision to move his finger. W clearly isn't the cause of the finger-movement. These are the facts.

One can always play skeptic. "Libet's experimental data is fraudulent." "The actions studied aren't real actions." "The actions studied are real actions but they aren't representative of actions in general." "The relation between RP and finger-movement is epiphenomenal, not causal." These are cards that can be played, but in this context that is all they are. Judging by the various replication-studies and follow-up studies, there is no reason to believe that Libet's data is in any way spurious and there is every reason to believe that RLE generalize.

Consider the standards of proof and evidence that philosophers are using in connection with the data generated by the Libet experiment. By those standards, there is no reason to believe that smoking causes cancer. But it does, and we know that, because that is what the data shows.

If it turns out that the data generated by Libet's experiment and by the various replication and follow-up studies is spurious, so be it. In the meantime, we should probably accept that data and try to model. By refusing to model it—by insisting that it mut be spurious---philosophers are, so we have reason to believe, blinding themselves to valuable information bout the nature of agency, specifically of our so-called "conscious decisions" therein.

We have put forth a model, albeit an extremely schematic one, as to the agentive function of such mental events. If nothing else, our model is consistent with the data generated by Libet's experiment. We stand to learn about the mind by investigating the merits of such models. We do not stand to do so by insisting that Libet "must" be wrong. I regard Libet's experiment as an extremely valuable but missed opportunity for the discipline of philosophy, and the present paper was intended to be an attempt, however jejune, to make good on that opportunity.

References

Batthyany, A. (2009). Mental causation and free will after Libet and Soon: reclaiming conscious agency.

Bayne, T. (2011). Libet and the case for free will scepticism.

Bernáth, L. (2019). Why libet-style experiments cannot refute all forms of libertarianism. In *Free will, causality, and neuroscience* (pp. 97-119). Brill.

Bonicalzi, S. (2015). Libet-like experiments and the efficacy of the will. *Rivista internazionale di Filosofia e Psicologia*, *6*(1), 130-144.

Braun, M. N., Wessler, J., & Friese, M. (2021). A meta-analysis of Libet-style experiments. *Neuroscience & Biobehavioral Reviews*, *128*, 182-198.

Dennett, D. C. (2015). *Elbow Room, new edition: The Varieties of Free Will Worth Wanting*. mit Press.

Fischborn, M. (2016). Libet-style experiments, neuroscience, and libertarian free will. *Philosophical Psychology*, *29*(4), 494-502.

Fosu-Blankson, F., & Inusah, H. (2023). A Critique of Libet and Wegner's Argument Against Free Will. *Journal of NeuroPhilosophy*, *2*(1).

Freud, S. (2005). *The unconscious* (Vol. 8). Penguin UK.

Fried, I., Mukamel, R., & Kreiman, G. (2011). Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. Neuron, 69(3), 548-562.

Griffith, M. (2013). Free Will and Modern Science. *Analysis*, *73*(1), 168-176.

Haggard, P. (2011). Decision time for free will. *Neuron*, *69*(3), 404-406.

Haggard, P., & Eimer, M. (1999). On the relation between brain potentials and the awareness of voluntary movements. Experimental Brain Research, 126(1), 128-133.

Knowles, A. Free Will Stands When Properly Explained and Correctly Defined and Neuroscience Shows This to Be the Case.

Libet, B (1993). The neural time—factor in perception, volition and free will. *Neurophysiology of Consciousness*, 367-384.

Libet, B. (1999). Do we have free will?. *Journal of consciousness studies*, *6*(8-9), 47-57.

Libet, B. (2002). The timing of mental events: Libet's experimental findings and their implications. *Consciousness and cognition*, *11*(2), 291-299.

Libet, B. (2003). Can conscious experience affect brain activity?. *Journal of consciousness studies*, *10*(12), 24-28.

Libet, B. (2022). Mind time. In *Mind Time*. Harvard University Press.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). Brain, 106(3), 623-642

Libet, B., Wright Jr, E. W., & Gleason, C. A. (1983). Preparation-or intention-to-act, in relation to pre-event potentials recorded at the vertex. *Electroencephalography and clinical Neurophysiology*, *56*(4), 367-372.

Løvstad, M., Funderud, I., Meling, T., Krämer, U. M., Voytek, B., Due-Tønnessen, P., ... & Solbakk, A. K. (2012). Anterior cingulate cortex and cognitive control: neuropsychological and electrophysiological findings in two patients with lesions to dorsomedial prefrontal cortex. *Brain and cognition*, *80*(2), 237-249.

Mele, A. R. (2009). *Effective intentions: The power of conscious will*. Oxford University Press on Demand.

Mele, A. R. (2010). Libet on free will: Readiness potentials, decisions, and awareness. *Conscious will and responsibility: A tribute to Benjamin Libet*, 23-33.

Mele, A. (2019). Free will and neuroscience: decision times and the point of no return. In *Free will, causality, and neuroscience* (pp. 83-96). Brill.

Pollen, D. A. (2004). Brain stimulation and conscious experience. *Consciousness and Cognition*, *13*(3), 626-645.

Pollen, D. A. (2006). Brain stimulation and conscious experience: Electrical stimulation of the cortical surface at a threshold current evokes sustained neuronal activity only after a prolonged latency. *Consciousness and Cognition*, *15*(3), 560-565.

Radder, H., & Meynen, G. (2013). Does the brain "initiate" freely willed processes? A philosophy of science critique of Libet-type experiments and their interpretation. *Theory & Psychology*, *23*(1), 3-21.

Rescher, N. (2014). Evidentiating Free Will. *The Journal of Speculative Philosophy*, *28*(1), 79-106.

Roskies, A. L. (2010). How does the neuroscience of decision-making bear on our understanding of moral responsibility and free will? Current Opinion in Neurobiology, 20(6), 1-5.

Schlegel, A., Alexander, P., Sinnott-Armstrong, W., Roskies, A., Tse, P. U., & Wheatley, T. (2015). Hypnotizing Libet: Readiness potentials with non-conscious volition. *Consciousness and Cognition*, *33*, 196-203.

Schlosser, M. E. (2014). The neuroscientific study of free will: A diagnosis of the controversy. *Synthese*, *191*(2), 245-262.

Seifert, J. (2011). In defense of free will: a critique of Benjamin Libet. *The Review of Metaphysics*, 377-407.

Soon, C. S., Brass, M., Heinze, H. J., & Haynes, J. D. (2008). Unconscious determinants of free decisions in the human brain. Nature Neuroscience, 11(5), 543-545.

Watson, G. (1987). Free action and free will. *Mind*, *96*(382), 145-172.

Wegner, D. M. (2004). Précis of the illusion of conscious will. *Behavioral and Brain Sciences*, *27*(5), 649-659.